Exact analytic solution of the multi-dimensional Anderson localization

V N Kuzovkov†‡, W von Niessen‡

† Institute of Solid State Physics, University of Latvia, 8 Kengaraga Street, LV – 1063 RIGA, Latvia

‡ Institut für Physikalische und Theoretische Chemie, Technische Universität Braunschweig, Hans-Sommer-Straße 10, 38106 Braunschweig, Germany

E-mail: kuzovkov@latnet.lv

Abstract. The method proposed by the present authors to deal analytically with the problem of Anderson localization via disorder [J.Phys.: Condens. Matter **14** (2002) 13777] is generalized for higher spatial dimensions D. In this way the generalized Lyapunov exponents for diagonal correlators of the wave function, $\langle\psi^2_{n,\mathbf{m}}\rangle$, can be calculated analytically and exactly. This permits to determine the phase diagram of the system. For all dimensions $D > 2$ one finds intervals in the energy and the disorder where extended and localized states coexist: the metal-insulator transition should thus be interpreted as a first-order transition. The qualitative differences permit to group the systems into two classes: low-dimensional systems ($2 \leq D \leq 3$), where localized states are always exponentially localized and high-dimensional systems ($D \geq D_c = 4$), where states with non-exponential localization are also formed. The value of the upper critical dimension is found to be $D_0 = 6$ for the Anderson localization problem; this value is also characteristic of a related problem - percolation. Consequences for numerical scaling and other approaches are discussed in detail.

Submitted to: *J. Phys.: Condens. Matter*

PACS numbers: 72.15.Rn, 71.30.+h

1. Introduction

1.1. Experiment and theory

Disorder leads to important physical effects which are of quantum mechanical origin. This has been revealed by Anderson [1] in the study of a disordered tight-binding model. This problem has attracted great attention over many decades. A breakthrough came with the scaling theory of localization [2]. All states in a one-dimensional system (1-D) are localized, whereas in 3-D a metal-insulator transition occurs. An analytic solution is only known for the 1-D problem [3]. Although there was no general analytical solution available, there was consensus that in 2-D all states are localized.

For quite some time after the advent of the scaling theory, many believed it to be essentially under control. This view is less secure nowadays, in part because recent experiments have challenged conventional wisdom about disordered 2-D systems. The 2-D case still presents a problem which has become apparent by experiments [4, 5, 6, 7]. These experiments are still being discussed controversially. Experimental reality is certainly more complex than a simple tight-binding model, but these results provide a good reason for revisiting this classic theoretical problem.

Recently we have been able to solve the 2-D case analytically [8]. We have shown that in principle there is the possibility that the phase of delocalized states exists for a *non-interacting electron system*. For energies and disorder, where extended states may exist we find a coexistence of these localized and extended states. Thus the Anderson metal-insulator transition exists and should be regarded as a first order phase transition. Consequently we have returned to the old idea of Mott [9, 10] that the metal-insulator transition is discontinuous. This alternative idea was in its history completely abandoned with the advent of the scaling theory of localization. There is now a renaissance of it.

This result implies the failure of the scaling theory of Anderson localization. Although this paper is published [8] and constitutes the basis for a new analytical investigation of the Anderson problem in the present paper for the case of higher dimensions (N-D problem) one has to accept the following: (i) the paper [8] requires an independent confirmation, which requires a certain time; (ii) in the history of the problem one has developed quite a few conceptions and this leads to a critical attitude towards the new theory - i.e. there is a resistance - as it departs from conventional wisdom. This asks for a critical evaluation of the new theory and its results. One should acknowledge that the problem is many-sided and quite complex. Thus logical errors are possible which do not lie at the surface. Perhaps a reference to B. Pascal is appropriate here: *A truth is so delicate that any small deviation from it leads you to a mistake, but this mistake is also so delicate that after a small retreat you find yourselves in a truth again.* Without an exact analytic solution any discussion cannot lead to firm results. The main aim of the present article is thus the generalization of the mathematical tools of the previous article[8] to the case of higher dimensional spaces and a physical interpretation of the new results.

1.2. Structure of the present article

The outline of the present article is a follows. In chapter 2 we give a short derivation of equations for N-D Anderson localization problem, and necessary summary of the results of the article [8]. The connection between the Anderson localization problem

and signal theory is discussed and the most important concept of the proposed methode - the filter $H(z)$ - is defined. To understand fully these aspects of the theory a knowledge of the first paper [8] is recommended. The filter $H(z)$ is generalized for higher dimensions D of the space. The investigation of its properties and the corresponding physical interpretation constitute the content of several chapters. The theory for high-dimensional systems ($D \geq 4$) is presented in chapter 3, whereas the theory for low-dimensional systems is given in chapter 4. An appendix deals with the mathematical conditions for the physical interpretation of the filter $H(z)$ and a more detailed discussion of related aspects of the problem.

2. New methods

2.1. Equations for correlators

Recently we have been able to solve the 2-D case analytically [8]. The tight-binding equation in 2-D is solved for the wave function $\psi_{n,m}$ and the second moments (correlators) $\langle\psi^2_{n,m}\rangle$:

$$\psi_{n+1,m} = (E - \varepsilon_{n,m})\psi_{n,m} - \psi_{n-1,m} - \sum_{m'} \psi_{n,m+m'}, \tag{1}$$

where the summation over m' runs over the nearest neighbours of site (n, m) in layer n that is in a space of dimension $p = D-1 = 1$. We assume taking the limit to infinite size $L \to \infty$ in p-space. The equation is solved with an initial condition

$$\psi_{0,m} = 0, \psi_{1,m} = \alpha_m. \tag{2}$$

The on-site potentials $\varepsilon_{n,m}$ are independently and identically distributed with existing first two moments, $\langle\varepsilon_{n,m}\rangle = 0$ and $\left\langle\varepsilon^2_{n,m}\right\rangle = \sigma^2$. After averaging the mean squared amplitude $\langle\psi^2_{n,m}\rangle = x_n$ becomes independent of m (the averaging procedure that justifies this statement is discussed in detail in [8]).

In the present paper we present the analytic solution for the general case $D > 2$. The knowledge of paper [8] is prerequisite for understanding the present one as it contains the full formalism. The generalization of this formalism to the N-D case presents no problem (see below). The main eqs. (1),(2) remain valid, only scalar quantities become vector quantities.

The tight-binding equation in $1 + p$ dimension is (primed indices are summed)

$$\psi_{n+1,\mathbf{m}} = -\varepsilon_{n,\mathbf{m}}\psi_{n,\mathbf{m}} - \psi_{n-1,\mathbf{m}} + \mathcal{L}_{\mathbf{m},\mathbf{m}'}\psi_{n,\mathbf{m}'}, \tag{3}$$

$$\mathcal{L}_{\mathbf{m},\mathbf{m}'} = E\delta_{\mathbf{m},\mathbf{m}'} - \sum_{\mathbf{m}''} \delta_{\mathbf{m}+\mathbf{m}'',\mathbf{m}'}, \tag{4}$$

(summation over $\mathbf{m}''$ runs over the nearest neighbours) with initial condition $\psi_{0,\mathbf{m}} = 0$ and $\psi_{1,\mathbf{m}} = \alpha_{\mathbf{m}}$.

One introduces the correlators (the averages are taken over disorder)

$$x(n)_{\mathbf{m},\mathbf{l}} = \langle\psi_{n,\mathbf{m}}\psi_{n,\mathbf{l}}\rangle\,, \tag{5}$$

$$y(n)_{\mathbf{m},\mathbf{l}} = \langle\psi_{n,\mathbf{m}}\psi_{n-1,\mathbf{l}}\rangle\,. \tag{6}$$

From the eq.(3) one easily obtains the system of equations:

$$x(n+1)_{\mathbf{m},\mathbf{l}} = \delta_{\mathbf{m},\mathbf{l}}\sigma^2 x(n)_{\mathbf{m},\mathbf{l}} + x(n-1)_{\mathbf{m},\mathbf{l}} + \tag{7}$$
$$\mathcal{L}_{\mathbf{m},\mathbf{m}'}x(n)_{\mathbf{m}',\mathbf{l}'}\mathcal{L}_{\mathbf{l}',\mathbf{l}} - \mathcal{L}_{\mathbf{m},\mathbf{m}'}y(n)_{\mathbf{m}',\mathbf{l}} - \mathcal{L}_{\mathbf{l},\mathbf{l}'}y(n)_{\mathbf{l}',\mathbf{m}},$$

$$y(n+1)_{\mathbf{m},\mathbf{l}} = -y(n)_{\mathbf{l},\mathbf{m}} + \mathcal{L}_{\mathbf{m},\mathbf{m}'}x(n)_{\mathbf{m}',\mathbf{l}}. \tag{8}$$

They can be solved explicitly by introducing of the Z-transform[11]

$$X(z)_{\mathbf{m},\mathbf{l}} = \sum_{n=0}^{\infty} \frac{x(n)_{\mathbf{m},\mathbf{l}}}{z^n}, \tag{9}$$

$$Y(z)_{\mathbf{m},\mathbf{l}} = \sum_{n=0}^{\infty} \frac{y(n)_{\mathbf{m},\mathbf{l}}}{z^n}, \tag{10}$$

which turn the equations into

$$(z - z^{-1} - \sigma^2\delta_{\mathbf{m},\mathbf{l}})X(z)_{\mathbf{m},\mathbf{l}} - x(1)_{\mathbf{m},\mathbf{l}} = \mathcal{L}_{\mathbf{m},\mathbf{m}'}X(z)_{\mathbf{m}',\mathbf{l}'}\mathcal{L}_{\mathbf{l}',\mathbf{l}} \tag{11}$$

$$-\mathcal{L}_{\mathbf{m},\mathbf{m}'}Y(z)_{\mathbf{m}',\mathbf{l}} - \mathcal{L}_{\mathbf{l},\mathbf{l}'}Y(z)_{\mathbf{l}',\mathbf{m}},$$

$$zY(z)_{\mathbf{m},\mathbf{l}} = -Y(z)_{\mathbf{l},\mathbf{m}} + \mathcal{L}_{\mathbf{m},\mathbf{m}'}X(z)_{\mathbf{m}',\mathbf{l}}. \tag{12}$$

Iteration of the second equation yields

$$Y(z)_{\mathbf{m},\mathbf{l}} = \frac{z}{z^2-1}\mathcal{L}_{\mathbf{m},\mathbf{m}'}X(z)_{\mathbf{m}',\mathbf{l}} - \frac{1}{z^2-1}X(z)_{\mathbf{m},\mathbf{l}'}\mathcal{L}_{\mathbf{l}',\mathbf{l}}. \tag{13}$$

We then obtain an equation for $X(z)_{\mathbf{m},\mathbf{l}}$ only, which can be solved by double Fourier expansion. However, upon averaging over ensemble of initial conditions (as described in Section 3.2 of [8]): averaging over translations in p-space of boundary conditions, $\alpha_{\mathbf{m}}$) such that $\overline{\alpha_{\mathbf{m}}\alpha_{\mathbf{m}'}} = \Gamma_{\mathbf{m}-\mathbf{m}'}$, the problem is translation-invariant in transverse directions. We than put:

$$X(z)_{\mathbf{m},\mathbf{l}} = \int \frac{d^p\mathbf{k}}{(2\pi)^p}X(z,\mathbf{k})e^{i\mathbf{k}(\mathbf{m}-\mathbf{l})}. \tag{14}$$

After averaging on initial conditions the diagonal correlator becomes independent of $\mathbf{m}$:

$$x(n)_{\mathbf{m},\mathbf{m}} \equiv x_n, \tag{15}$$

$$X(z)_{\mathbf{m},\mathbf{m}} \equiv X(z) = \int \frac{d^p\mathbf{k}}{(2\pi)^p}X(z,\mathbf{k}). \tag{16}$$

We obtain the final equations

$$\frac{(z-1)}{(z+1)}[w^2 - \mathcal{E}^2(\mathbf{k})]X(z,\mathbf{k}) = \Gamma(\mathbf{k}) + \sigma^2 X(z), \tag{17}$$

$$\mathcal{E}(\mathbf{k}) = E - 2\sum_{j=1}^{p}\cos(k_j), \tag{18}$$

$$w^2 = \frac{(z+1)^2}{z}, \tag{19}$$

or

$$X(z) = H(z)X^{(0)}(z), \tag{20}$$

$$X^{(0)}(z) = \frac{(z+1)}{(z-1)}\int \frac{d^p\mathbf{k}}{(2\pi)^p}\frac{\Gamma(\mathbf{k})}{[w^2 - \mathcal{E}^2(\mathbf{k})]}, \tag{21}$$

$$\frac{1}{H(z)} = 1 - \sigma^2\frac{(z+1)}{(z-1)}\int \frac{d^p\mathbf{k}}{(2\pi)^p}\frac{1}{[w^2 - \mathcal{E}^2(\mathbf{k})]}, \tag{22}$$

where the $X^{(0)}(z)$ (or $x_n^{(0)}$) refer to the ordered system ($\sigma \equiv 0$) and the $X(z)$ (or x_n) to the disordered one ($\sigma \neq 0$). For eq.(20) the inverse Z-transform gives convolution property [11]:

$$x_n = \sum_{l=0}^{n} x_l^{(0)}h_{n-l}, \tag{23}$$

with

$$H(z) = \sum_{n=0}^{\infty} \frac{h_n}{z^n}. \tag{24}$$

2.2. Anderson localization and signal theory

The essential point in the analysis with respect to the localized or extended character of the states is to make use of signal theory [11] from electrical engineering and switch from an investigation of the moments x_n to an analysis of the filter functions h_n.

In the theory of signals [11], $x_n^{(0)}$ (or $X^{(0)}(z)$) is the input signal, and x_n (or $X(z)$) is the output. Asymptotic behaviour of the solution is completely determined by the filter h_n (or $H(z)$). The concept of the system function is a general and abstract description of the problem of localization. Thus the filter function has to be analysed to obtain general results and not the multitude of signals. This has been done in [8] for the 2-D case. This procedure has a certain similarity with the transition to an operator formalism in quantum mechanics. Particular signals x_n depend on the initial conditions used and do not carry much physical information because of the unconventional normalization [8]. For the localization problem the only property that matters is whether a signal belongs to the bounded or unbounded class and this can be derived from the filter.

The essence of localization is contained in the filter $H(z)$. We study the filter $H(z)$ with properties described by generalized Lyapunov exponents. The filter is a fundamental function of the disorder σ only [8].

A filter h_n is uniquely characterized by a pole-zero diagram of its image $H(z)$ which is a plot of the locations of the poles λ_i and zeros of $H(z)$ in the complex-z plane. We provide just a brief summary here, for more details consult [8, 11]. The signals $x_n^{(0)}$ and x_n are real, therefore $H(z)$ will have poles and zeros that are either on the real axis, or come in conjugate pairs. For the inverse Z-transform $H(z) \Rightarrow h_n$ one needs to know the *region of convergence* (ROC). Physical considerations dictate that only *causal* filters ($h_n = 0$ for $n < 0$) should be considered. They have ROCs outside a circle that intersects the pole with $\max|\lambda_i|$. A causal filter is *stable* (bounded input yields a bounded output) if the *unit circle* $|z| = 1$ is in the ROC. Note that the explicit calculation of h_n by the inverse Z-transform is not necessary, and it is also not feasible analytically due to the complexity of the function $H(z)$. Only the type of the filter – stable or unstable – needs to be determined. The delocalized states (bounded output) are obtained by transforming the physical solutions inside the band (bounded input) provided that the filter $H(z)$ is stable. Seeking for poles is quite a simple analytical task which gives rather general results by elementary methods.

As an example for the general and abstract description of the problem of localization as stated above let us consider the following problem. It is well-known that disorders broadens the band; new states outside the old band arise for $|E| > E_b = 2D$. Are among these new states also extended states? Numerical work for $D = 3$ [12, 13] ascertains this. This is the socalled reentrant behaviour of the mobility edge: the change from localized to extended states and back to localized ones upon increasing the disorder occurs for certain fixed energies. Because in the literature there is no physical explanation for this phenomenon one has simply accepted these results without critically examining them.

With the help of signal theory we have found a particular transformation, which

gives a connection between the states in an ideal system (zero disorder) and states in a system with disorder. Extended states (zero disorder) as input (bounded) signal transform into localized states (nonzero disorder) as output (unbounded) signal if the filter which is responsible for this transformation is unstable. If the filter is stable, then extended states (bounded input signal) transform into extended states (bounded output signal). It is known that for zero disorder i.e. outside the band, $|E| > E_b$, there do exist only mathematical solutions which cannot be normalized. These correspond to an unbounded input signal. It is impossible to find a filter which permits a transformation of the type unbounded input signal (mathematical solution) into a bounded output signal (extended states in $|E| > E_b$). The reverse - the transformation of an unbounded input signal (mathematical solution) into a unbounded output signal (localized states in $|E| > E_b$) - is on the other hand possible. I.e. the mathematical procedure developed by us generates in this case a negative answer to the posed question. Because this result contradicts the numerical work [12, 13], it is necessary to discuss in detail the quality of the numerical work (see below and Appendix B).

We have shown in [8] that the filter $H(z)$ is a non-analytic function of the complex variable z; this result remains valid also in the multi-dimensional case. The unit circle $|z| = 1$ divides the complex plane into two analytic domains: the interior and exterior of the unit circle. The inverse Z-transform is quite generally defined via countour integrals in the complex plane

$$h_n = \frac{1}{2\pi i} \oint H(z) z^n \frac{dz}{z}. \tag{25}$$

and this definition is only possible in an analytic domain. In this way in the formal analysis of the problem multiple solutions result. The first solution $H_+(z)$ is defined outside the unit circle and always exists. The filter $H_+(z)$ describes localized states and it is possible to connect its properties with the notion of the localization length [8]. The second solution $H_-(z)$ is defined inside the unit circle and does not always represent a solution which can be physically interpreted (this is the mathematical consequence that the filter be causal). The filter $H_-(z)$ describes delocalized states. The coexistence of the two solutions was physically interpreted in [8] as the coexistence of two phases – an insulating and a metallic one. Then the metal-insulator transition should be looked at from the basis of first-order phase transition theory.

2.3. Conformal mapping

The p-dimensional integral on the r.h.s. of eq. (22) can be reduced to a one-dimensional integral. Consider the identity

$$\frac{1}{w^2 - \mathcal{E}^2(\mathbf{k})} = \int_{-\infty}^{\infty} \frac{\delta\left(y + \sum_{j=1}^{p} 2\cos(k_j)\right) dy}{w^2 - (E+y)^2}. \tag{26}$$

The integral representation of the Dirac δ-function and the Bessel function

$$J_0(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ix\cos(k)} dk. \tag{27}$$

will be used. Following [8], let us define a complex parameter $w = u + iv$ in the upper half-plane, $v = Im(w) \geq 0$. Using the methods of complex variable theory we get

$$\frac{1}{(2\pi)^p} \int \frac{d\mathbf{k}}{w^2 - \mathcal{E}^2(\mathbf{k})} = \frac{1}{iw} Y_D(w, E), \tag{28}$$

where

$$Y_D(w,E) = \int_0^\infty [J_0(2t)]^{D-1} \cos(Et) \exp(iwt)\, dt. \tag{29}$$

Changing the complex variable z to the parameter w corresponds to the conformal mapping of the inner part [$|z| \leq 1$, $w = -(z^{1/2} + z^{-1/2})$)] or the outer part [$|z| \geq 1$, $w = (z^{1/2} + z^{-1/2})$)] of the circle onto the upper half-plane. The circle itself maps onto the interval $[-2, 2]$. Note also that if $H(z) = 0$ has complex conjugate poles, then on the upper w half-plane they differ only by the sign of $u = Re(w)$. To avoid complicated notations, we seek for poles in the sector $u \geq 0, v \geq 0$ and double their number if we find any.

The inverse function [8]

$$z = -1 + \frac{w^2}{2} \pm \frac{w}{2}\sqrt{w^2 - 4} \tag{30}$$

is double-valued. Its branch with the minus sign maps the w sector onto the inner part of the half-circle ($|z| \leq 1$). The second branch with the plus sign gives a mapping onto the half-plane with the half-circle excluded ($|z| \geq 1$). Therefore in the parametric w-representation

$$\frac{(z+1)}{(z-1)} = \pm \frac{w}{\sqrt{w^2 - 4}} \tag{31}$$

and

$$\frac{1}{H_\pm(z)} = 1 \pm \sigma^2 i \frac{Y_D(w,E)}{\sqrt{w^2 - 4}}. \tag{32}$$

3. High-dimensional systems

3.1. Upper critical dimension

It is generally assumed that 2-D systems mark the borderline between high and low dimension [14]. The existence of a transition in 3-D is not questioned (high-dimensional systems). These assumptions, however, originate from the scaling theory of localization. Here the marginal dimension is $D_M = 2$, and a phase transition exists only for $D > D_M$. Thus perturbation theory [4] for $D = 2 + \varepsilon$ ($\varepsilon \ll 1$) is possible. The effect of statistical fluctuations cause a change of regime at $D_c = 4$ [10, 15]; in this way the upper critical dimension for localization D_c arises. For $D > D_c$ there should not exist a phase transition.

On the one hand these statements referring to higher dimensions are numerically nearly impossible to ascertain [16]. Statistics is bad and the length of the system L very small (see Appendix B). Even for $D = 3$ progress towards extracting reliable numerical estimates of critical quantities has been remarkably difficult [17]. On the other hand, if the results from the scaling theory of localization for 2-D systems are faulty (this is what we claim), then the corresponding division of the systems into low and high dimensional ones is also wrong. Here one must develop an alternative picture.

In the theory of critical phenomena [18, 19] many systems belong to a class, where an upper critical dimension D_0 has a totally different physical meaning. It denotes the dimension, from where on the mean field approximation is exact, or where with other words all critical exponents reach stationary values. I.e. in this

case the phase transition does exist also for $D > D_0$; only in the limit $D \to \infty$ the transition disappears, simply because the corresponding critical values have gone to infinity. Systems of quite different physical nature may have the same value of the upper critical dimension D_0. E.g. one finds $D_0 = 4$ not only for the theory of magnetism [18, 19] (in this case there exists also the marginal dimension, below which no phase transition is possible), but also in kinetics (cooperative phenomena in bimolecular processes by diffusion-controlled reactions) [20, 21]; in the latter case, however, there is no marginal dimension. Another example is percolation, where the upper dimension is $D_0 = 6$ [22].

3.2. Stability and poles: solution $H_+(z)$

Let us start first from a purely mathematical comment: the integral $Y_D(w, E)$ is always finite for all $w = u + iv$ in the sector $u \geq 0, v \geq 0$ for high-dimensional systems with $D \geq 4$. This fact clearly follows from the asymptotic behaviour of the Bessel function for large values of its argument, $J_0(2t) \approx \frac{1}{\sqrt{\pi t}} \cos(2t - \pi/4)$. We shall further on see that the dimensionality $D = D_c = 4$ is critical for localization, although this is no proper upper critical dimension $D = D_0$, which we shall determine below. Let us consider therefore the properties of the filter-functions in this region of the value of the dimension of space.

Let us consider first the solution $H_+(z)$. According to [11], the ROC of the causal filter is defined by the inequality $|z| > \max|\lambda_i|$, where λ_i are the poles. The case when the system function has a pole at $z = \lambda > 1$ is the simplest one for an interpretation. In terms of signal theory [11] the filter $H_+(z)$ is unstable since the pole lies outside the unit circle $|z| = 1$ in the complex z-plane. As shown in [8], such a filter describes exponentially localized states (insulating phase). In order to see this one can exploit a basic inverse Z-transform [11]:

$$H(z) = z/(z - \lambda) \Rightarrow h_n = \lambda^n. \tag{33}$$

Therefore the pole of $H_+(z)$ at $z = \lambda = \exp(2\gamma)$ leads to exponential growth of the system function h_n which in turn implies [eq. (23)] an exponentially increasing mean squared amplitude x_n. The growth exponent γ is the so-called generalized Lyapunov exponent [3] related to the localization length by $\xi = \gamma^{-1}$.

The value of the Lyapunov exponent γ defines the phase. We start from the mathematical definition that all states with $\gamma \neq 0$ belong to the insulating phase. The states with $\gamma \equiv 0$ on the other hand correspond to a metal. According to this definition the states with non-exponential localization also belong to the metallic phase, because they correspond to the value $\gamma \equiv 0$. We can consider these states as a bad metal, in contrast to a good metal, where one has truly delocalized states.

We would like to give here a summary of the results which emerge from an analysis of the pole diagram (for details see the Appendix A). The function $Y_D(w, E)$ defined by eq. (29) is purely imaginary for $v = 0$, $u > u_0$,

$$u_0 = 2p + |E|, \tag{34}$$

and the system function $H_+(z)$ itself is real. The pole must be located (if present at all) exactly in this region of the parameter w. It can be found as a solution of an in general transcendental equation

$$\sigma^2 \Omega_+(u) = 1, \tag{35}$$

where

$$\Omega_+(u) = \frac{1}{\sqrt{u^2-4}} \int_0^\infty [J_0(2t)]^{D-1} \cos(Et) \sin(ut) dt. \tag{36}$$

For high-dimensional systems with $D \geq 4$ the function $\Omega_+(u)$ has a maximum at $u = u_0$ and decreases monotonically for $u > u_0$. The system function has a pole if the disorder exceeds a critical value, $\sigma > \sigma_0(E)$, where

$$\sigma_0(E) = \Omega_+(u_0)^{-1/2}. \tag{37}$$

Therefore, in high-dimensional systems exponential localization takes place only if the disorder is strong enough.

For $\sigma < \sigma_0(E)$ the function $H_+(z)$ has no poles, its ROC its $|z| \geq 1$. This means that the unit circle $|z| = 1$ belongs to the ROC, and the filter is stable. We interpret solutions for this range of the disorder values, σ, also as localized, however with non-exponential localization (which could be a power-law). Here we encounter limits of applicability of the method, which come into play, however, only in physically inaccessible systems of high dimensionality. The elementary pole search can be applied only for exponentially localized states, the general case requires a detailed investigation of the filter.

The curve for $\sigma_0(E)$ (Fig.1a) is the well-known mobility edge. For $\sigma > \sigma_0(E)$ there exists an insulating phase, whereas for $\sigma < \sigma_0(E)$ a metallic phase is found (bad metal). A coexistence of phases is here not possible. Thus the phase transition here has an appearance as if it were a transition of second order. This simple idea, however, is contradicted by the behaviour of the Lyapunov exponent: the transition from $\gamma \equiv 0$ to $\gamma \neq 0$ is not continuous. One can clearly see that all these mobility edges for high-dimensional systems show the same qualitative behaviour.

3.3. Stability and poles: solution $H_-(z)$

Now let us turn to the second branch of the solution, $H_-(z)$. In this case existence of the poles leads to principally different consequences. Let us assume that the corresponding value of the parameter w is found and the pole $z = \lambda_1$ is located inside the unit circle, $|\lambda_1| = 1/\lambda$ with $\lambda > 1$. Formally, however, from the definition $w = -(z^{1/2} + z^{-1/2})$ the same value of w can be obtained for $z = \lambda_2 = 1/\lambda_1 = \lambda$. The complex number λ_2 lies outside the region of definition of the solution, $|z| \leq 1$. In this sense the pole at $z = \lambda_2$ is virtual. For the inverse Z-transform this fact is, however, irrelevant. The ROC for a causal filter is defined by the inequality $|z| > \max|\lambda_i|$ or $|z| > \lambda > 1$. Since the ROC and the region of definition of the solution $|z| \leq 1$ do not intersect, a physical solution is absent. Therefore, the filter $H_-(z)$ as a physical solution is acceptable only if either there are no poles or they lie on the unit circle. The latter case is realized for $D = 2$ [8] and corresponds to so called marginal stability [11]. In the following, we consider the general case $D \geq 4$ from a unified point of view.

A pole of the first type is related to the singularity of the function eq. (32) at $w = 2$ (the root, for details see the Appendix A). We define the phase and amplitude via the integral (29)

$$Y_D(w, E) = I_D(w, E) \exp(i\vartheta_D(w, E)). \tag{38}$$

It is not difficult to show that this pole emerges at arbitrarily small disorder for a negative phase $\vartheta_D(2, E) < 0$. The equation $\vartheta_D(2, E_0) = 0$ defines the boundary of the region $|E| > E_0$, where the physical solution is absent and, therefore, any disorder

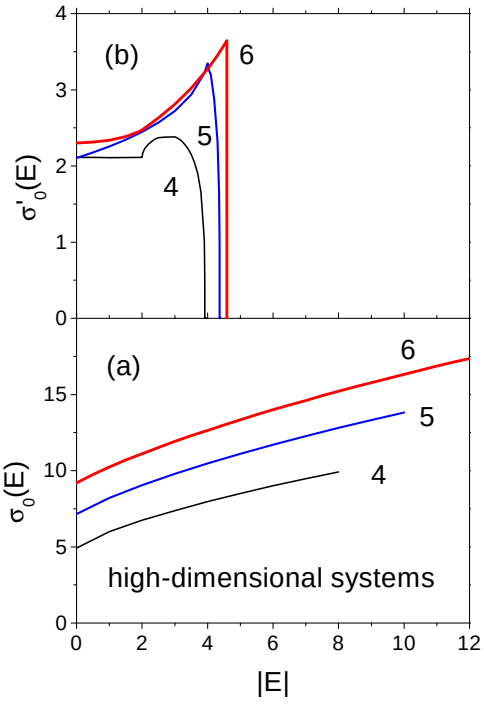

Figure 1. Threshold disorder values: (a) $\sigma_0(E)$ for the transition from the non-exponential to the exponential localization and (b) $\sigma_0'(E)$ for the transition from the delocalized to the (non-exponentially) localized states. The curves are enumerated with the values of D.

transforms the delocalized states into localized ones. For high-dimensional systems the delocalized states transform into states with non-exponential localization. The corresponding E_0 values are $E_0 = 3.915$ (D=4), $E_0 = 4.365$ (D=5) and $E_0 = 4.578$ (D=6).

For the region $|E| < E_0$ there exists the physical solution with $\sigma < \sigma_0'(E)$, where $\sigma_0'(E)$ is a second threshold disorder value. The behaviour of this curve $\sigma_0'(E)$ (Fig.1b) is determined by resonance phenomena. The integral in (29) consists asymptotically of a power function, $t^{(D-1)/2}$, and a product of trigonometric functions. E.g. the function $\cos(2t - \pi/4)$ comes from every Bessel function. If we represent this product as a sum of monochromatic waves, then we denote the existence of a wave with zero frequency as a resonance. For $w = 0$ the first non-trivial resonance lies either at $E_c = 2$ (D=4,6,...) or at $E_c = 4$ (D=5,7,...).

For $|E| < E_c$ a pole of the second type appearing at higher levels of disorder $\sigma > \sigma_0'(E)$ must be considered. This type of pole emerges at purely imaginary values

of the parameter $w = iv$. It corresponds to the roots of the equation

$$\sigma^2 \Omega_-(v) = 1, \tag{39}$$

where

$$\Omega_-(v) = \frac{1}{\sqrt{4+v^2}} \int_0^\infty [J_0(2t)]^{D-1} \cos(Et) \exp(-vt) dt. \tag{40}$$

A physical solution is acceptable here only if the poles lie on the unit circle, $|z| = 1$ or $w \in [0, 2]$ for our sector w. A marginally stable solution corresponds to the value $w = 0$ (or $v = 0$ in eq. (40)). The threshold disorder value is given by

$$\sigma_0'(E) = \Omega_-(0)^{-1/2}. \tag{41}$$

For $\sigma < \sigma_0'(E)$ the function $H_-(z)$ does not possess poles, here we find the region of stability of the extended states.

Spaces of dimension $D = 4, 5$ possess a certain sensitivity with respect to the resonance phenomena mentioned above. For $E_c < |E| < E_0$ the line of the poles exhibits a deviation from purely imaginary values of the parameter w and touches the real axis at the point $u' \in [0, 2]$. This point corresponds to the condition for the phase of the integral $\vartheta_D(u', E) = 0$ and also belongs to the unit circle (marginal stability). The corresponding threshold disorder value we denote again as $\sigma_0'(E)$:

$$\sigma_0'(E) = \left(\frac{\sqrt{4 - u'^2}}{I_D(u', E)} \right)^{1/2}. \tag{42}$$

We see that the function $\sigma_0'(E)$ is in general singular in the energy. In going from the energy value E_c to the value E_0 the parameter u' increases monotonically and reaches finally the value $u' = 2$. The function $\sigma_0'(E)$ goes continuously to zero for $|E| \to E_0$.

For dimensions $D \geq 6$ there are no resonance phenomena. These resonances are so weak that only one equation (41) remains valid in the whole range of energies $|E| < E_0$. I. e. although $D_c = 4$ is a certain critical dimension (here non-exponential localization arises), a qualitative agreement of all results is only obrtained for $D \geq 6$.

In this way it emerges from our exact analytic theory that for Anderson localization the upper critical dimension is $D_0 = 6$, i.e. the problem in a certain way shows a similarity to percolation [22] - and this may not be totally unexpected. However, we also note that as a rough estimate for the upper critical dimension the value $D_c = 4$ may also be accepted. It corresponds to a first and important step of the qualitative saturation of the results (the possibility of existence of non-exponentially localized states), whereas a complete saturation only obtains at D=6.

Quite generally in the range $|E| < E_0$ and under the condition $\sigma < \sigma_0'(E)$ there exists a region of existence of stable delocalized states. Because in this range also non-localized states coexist with the other ones, this means that in this range the good metal (delocalized states) and the bad metal (non-exponentially localized states) can form a heterogeneous system, whose properties depend on the relative proportions of the subsystems; e.g. the bad metal, if the subsystem of non-exponentially localized states percolates. Fig. 1 presents results of a numerical solution of the resulting equations.

4. Low-dimensional systems

4.1. Analytic solution

Independently from whether $D_c = 4$ or $D_0 = 6$ are taken as the upper critical dimension, it is necessary to consider 3-D systems as low dimensional ones. We approach here a widely held opinion that all states in $D = 3$ should be stable against perturbations and should have a finite radius of convergence. This, however, is not correct and would only be valid if the whole field of the scaling theory of localization would be faultless.

The eq.(35) has always a solution for the physically important cases $D = 2, 3$. We refer to these cases as 'low-dimensional' ones. For the low-dimensional systems the integral eq. (36) can be evaluated analytically (the formulas can be found in the tables of Laplace transforms). The corresponding pole diagrams are shown graphically in the Appendix A. The system function has a pole at $z = \lambda = \exp(2\gamma)$ with $\gamma > \gamma_0$, $2\cosh(\gamma_0) = u_0 = 2p + |E|$ and $p = D - 1$. Note, however, that the above mentioned feature of the low-dimensional systems is caused by the divergence of the integral $\Omega_+(u_0)$.

If one applies the obtained equations (41) for $\sigma_0(E)$ to low-dimensional systems with $D = 2, 3$, then formally $\sigma_0(E) \equiv 0$. Therefore, even infinitesimal disorder leads to solutions with exponential localization. Because the curve $\sigma_0(E)$ forms the border between exponentially localized and non-exponentialy localized states, this simply means that in this model for low-dimensional systems non-exponentially localized states are impossible (see, however, a discussion below).

For such systems the second curve $\sigma_0'(E)$ takes over the role of the mobility edge. The shape of this curve (Fig.2.b) for low-dimensional systems has a certain similarity with the ones for high-dimensional systems. Here we must stress that $\sigma_0'(E)$ does not represent a true mobility edge, but an upper limit for the coexistence of the two phases.

For $D = 2$ in [8] it has been found analytically that $\sigma_0'(E) = 2(1 - E^2/4)^{1/4}$ for $E < E_0$ where $E_0 = 2 = D$. I.e. in this case the value E_0 corresponds exactly to the midst of the band half-width $2D = 4$. For $E > E_0$ any disorder transforms the delocalized states into localized ones.

For $D = 3$ the value $E_0 = 3 = D$ corresponds again to the midst of the band half-width $2D = 6$. The resonance value mentioned above E_c lies for $D = 3$ at $E_c = 4$. Because $E_0 < E_c$, the resonance does not play a role here. The mobility edge $\sigma_0'(E)$ again follows from eq. (41).

The case of the 3-D system shows an exception in the energy range $E_0 < |E| < E_0' = 3.367$, where the function $\sigma_0'(E)$ exhibits so-called reentrant behaviour. The proof of this requires a special investigation (see pole diagrams in the Appendix A).

Note that for the 3-D system in the band center $\sigma_0(0) = 0$, because the integral in eq. (40) diverges. Fig. 2 presents results of a numerical solution of the resulting equations. These results are compared to the analytical result of [8] for the two-dimensional system.

4.2. Phase diagram and logarithmic divergence

Although the results for $D = 3$ are formally exact, they require a special discussion. The mathematically correct results are physically acceptable only if they are stable against small perturbations (e.g. small changes in the model definition, fluctuations of

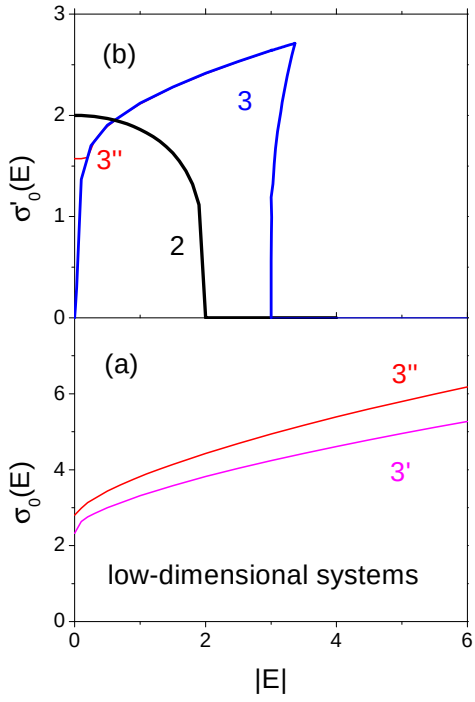

Figure 2. Threshold disorder values: (a) $\sigma_0(E)$ for the transition from the non-exponential to the exponential localization and (b) $\sigma_0'(E)$ for the transition from the delocalized to the localized states. The curves are enumerated with the values of D. The curves $3'$ and $3''$ correspond to the model eq.(43) for $\kappa = 0.01$ and $\kappa = 0.1$.

parameters in the model). From this point of view two results have to be questioned: (i) the singularity of $\sigma_0'(E)$ in the band center, $\sigma_0(0) = 0$, and (ii) the non-existence of non-localized states, $\sigma_0(E) \equiv 0$. Point (i) arises mathematically as a consequence of the logarithmic divergence of the integral (40) for $v = 0$ and $|E| \to 0$. Point (ii) arises again via a logarithmic divergence of the integral (36) for $u = u_0$. I.e. basically, the case $D = 3$ is nothing else but a type of logarithmic deviation from the high-dimensional case. Under certain conditions one can find the perturbations which are capable of regularizing the mentioned logarithmic divergence, i.e. to transform them into a finite term. One could e.g. surmise that in this regularization correlated disorder [23] might play an important role. One must stress here that we are dealing only with results for tight-binding Hamiltonians with diagonal disorder. It is largely unclear, whether this property remains valid also for non-diagonal disorder.

To illustrate this let us first consider a purely mathematical problem. Is it at all possible to confirm the results (i) and (ii) either by different numerical computations

or analytical approaches? The answer is - no! We have pointed out (Appendix B) that any deviation from the exact solution (via numerical or analytical ways) automatically generates results of a mean-field theory. A mean-field theory has a certain qualitative agreement with the results of the exact theory, but only for high-dimensional systems. Everything looks as if any uncontrolled deviation from the exact theory has added additional dimensions to a 3-D system. As a mathematical model let us further consider a 4-D system, where, however, a coupling involving this additional dimension is exceedingly week (parameter $\kappa \ll 1$). We start from a generalization of the function (18) for $D = 3$:

$$\mathcal{E}(\mathbf{k}) = E - 2\sum_{j=1}^{2} \cos(k_j) - 2\kappa \cos(k_3). \tag{43}$$

Here the term involving κ corresponds to the hopping matrix element into the fourth dimension. Fig.2 gives the numerical results for 2 cases: $\kappa = 0.01$ (curve $3'$) and $\kappa = 0.1$ (curve $3''$). One clearly sees that all results, which have nothing to do with the logarithmic divergence, e.g. the entire curve $\sigma_0'(E)$ with the exception of the point $E = 0$, remain extremely stable. Even the change in the parameter E_0 lie are of the order of $O(\kappa^2)$.

The 'logarithmic' results on the other hand turn out to be completely unstable. Even the small values of the coupling (or regularization) parameter κ produce results which are in qualitative agreement with those for high-dimensional systems. The curves $3'$ and $3''$ look as if they were an extrapolation of the corresponding curves for $\sigma_0(E)$ from Fig. 1a. Here the inequality $\sigma_0(E) > \sigma_0'(E)$ also applies. Reducing the parameter κ slowly moves the mobility edge $\sigma_0(E)$ downwards, because the dependence on the parameter κ is extremely week (logarithmic), $\sigma_0(E) \sim 1/\ln(\kappa^{-1})$. Only for even smaller values of the parameter $\kappa \ll 0.01$ one might perhaps see traces of the exact results, because in this case one has $\sigma_0(E) < \sigma_0'(E)$ for $|E| < E_0$.

5. Conclusion

Although the formal investigation of the Anderson model of localization (tight-binding Hamiltonian with diagonal disorder) for higher spatial dimensions D might at first look very abstract, the exact analytical results supply us with clear physical consequences. The Anderson problem of localization and the percolation problem belong to the same class of critical phenomena: both have the same lower and upper critical dimensions. I.e., although the Anderson model appears to be much more complex and richer, certain fundamental results appear to be transferable. Percolation is possible for 2-D systems, this corresponds to the existence of a metal-insulator transition in disordered 2-D systems. In this sense there is no reason to believe that the existing contradictions between theory and experiment for 2-D systems point to an incompleteness of the Anderson model. On the contrary, our analytical investigation has shown that a tight-binding Hamiltonian is presumably sufficient for this purpose.

The main problem of the theory thus does not rest in the Hamiltonian, but rather in the interpretation of the results, which mainly derive from numerical work. Our analytical and exact results demonstrate the necessity of interpreting the phase transition in the framework of first order phase transition theory and this holds independently of the spatial dimension $D \geq 2$. If one, however, attempts to apply procedures which have only been developed for systems with a second order phase

transition (and this is the general case), one does not necessarily obtain wrong numbers, but an incomplete or even wrong interpretation. See an example in [8], where an analytical (exact) scaling function for the 2-D system has the same form as obtained by numerical scaling. The physical interpretation is, however totally different. Concerning this point we have supplied in the present article ample material for discussion (Appendix B). We hope that the necessary corrections of the numerical tools are possible to detect the first order phase transition.

The rather large value of the upper critical dimension for the Anderson localization (and the percolation) problem permits to consider 2-D and 3-D cases as low-dimensional systems. Thus a revisiting of the results is also necessary for 3-D systems. We have found that the 3-D case is nothing else but a type of logarithmic deviation from the high-dimensional case. As a consequence results a certain instability of the results, whose details are discussed in the text.

We give here a short summary of the main results.

- For the Anderson localization problem there exists an upper critical dimension $D_0 = 6$. This value is also characteristic of a related problem: percolation [22]. For $D \geq D_0$ all phase diagrams are qualitatively the same, only the corresponding critical values develop in a monotonic way. One can also say that this is the property of a mean-field theory, although in this case a mean-field theory does not exist as a closed theory.
- There exists also a second upper critical dimension $D_c = 4$, which has a different meaning. The states with non-exponential localization are formed only for $D \geq D_c$, whereas for $D < D_c$ localized states are always exponentially localized. This second upper critical dimension D_c divides the dimensions into two classes: high dimensions with $D \geq D_c$ and low dimensions with $D < D_c$.

The results for the nontrivial spatial dimensions $D > 1$ can be summarised as follows.

(i) All states with energies $|E| > E_0$ are localized at arbitrarily weak disorder. The value of E_0 depends on D and lies inside the band $E_0 < 2D$.

(ii) For $|E| < E_0$ states are only localized if the disorder σ exceeds a critical value $\sigma_0'(E)$, otherwise a two-phase system is formed from an insulating and a metallic one. This differs from the traditional point of view which considers the localization transition as a continuous (second order) transition. Should the standard interpretation of this system in the framework of first-order phase transition theory be applicable (which still has to be investigated) one can expect that qualitatively it has similar properties as other heterogeneous two-phase systems (e.g. the coexistence of water and ice). Then percolation problems might be important.

(iii) Exponential localization always exists for the physically important cases $D = 2, 3$. Non-exponential localization occurs only for higher dimensions $D \geq D_c = 4$. In this case for $|E| < E_0$ and $\sigma < \sigma_0'(E)$ first the heterogeneous system appears, where the difference between the two phases may be small (this has to be investigated). For $\sigma_0(E) > \sigma > \sigma_0'(E)$ one finds a homogeneous system with non-exponential localization. Only for $\sigma > \sigma_0(E)$ a system with exponential localization appears.

(iv) $\sigma_0'(E)$ is in general not an analytic function of the energy E. There exist certain resonances.

Acknowledgments

V.N.K. gratefully acknowledges the support of the Deutsche Forschungsgemeinschaft. This work was partly supported by the EC Excellence Centre of Advanced Material Research and Technology (contract N 1CA1-CT-2080-7007) and the Fonds der Chemischen Industrie. We also would like to thank a referee for an alternative formulation of section 2.1.

Appendix A. Pole diagrams

Parametric representation of the pole diagram

A filter h_n is characterized by a pole diagram of its image $H(z)$. The principal definition eq.(22), together with the two other ones, eqs. (28),(31), supply us with the parametric w-representation of the pole diagram, eq.(32). Let us rewrite this relation into the form

$$\frac{1}{H(z)} = 1 - \sigma^2 R(w, E), \tag{A.1}$$

where $R(w, E)$ is generally a complex function of the complex variable w and energy E. The main idea is quite simple. The function $H(z)$ has its poles where

$$\sigma^2 R(w, E) \equiv 1. \tag{A.2}$$

An elementary requirement for this is the condition on the argument of the complex function

$$\arg R(w, E) = 0, \tag{A.3}$$

because σ is positiv. For the given energy value E eq.(A.3) defines one or more lines (pole lines) in the complex w-plain. As we have already discussed in the text, for reasons of symmetry it suffices to analyse only a sector $w = u + iv$ with $u \geq 0$ and $v \geq 0$. Let us analyse one of these pole lines. For each point w on this line the position of the pole is determined via eq.(30). The corresponding value of the disorder σ is found from eq.(A.2):

$$\sigma = R(w, E)^{-1/2}. \tag{A.4}$$

Because for each value of w on the pole line there is associated a value of σ, it is possible to indicate by an arrow next to the line in the diagram in which direction the poles move with increasing disorder.

Filter $H_+(z)$

Let us consider first the simplest case, the filter $H(z) = H_+(z)$. Here the poles are connected with the notion of the Lyapunov exponent. In the parametric representation there exists a simple relation

$$w = 2\cosh(\gamma); \tag{A.5}$$

thus the eqs.(A.4),(A.5) together supply a connection between σ and γ.

For a 2-D system it is possible to analytically evaluate the corresponding function $R(w, E)$ [8]:

$$R(w, E) = \frac{1}{2\sqrt{w^2 - 4}}\left[\frac{1}{\sqrt{(w + E)^2 - 4}} + \frac{1}{\sqrt{(w - E)^2 - 4}}\right], \tag{A.6}$$

or

$$R(w,E) = \frac{1}{2\sqrt{w^2-4}} \times \qquad \text{(A.7)}$$
$$[\frac{1}{\sqrt{(w+E+2)(w+E-2)}} + \frac{1}{\sqrt{(w-E+2)(w-E-2)}}].$$

We clearly see that the function (A.7) possesses for the given energy several points, where it diverges. These values $w = -2-E$, $w = 2-E$, $w = E-2$ and $w = E+2$ are real. They are nothing else but the resonances discussed in the text. To detect these we investigate the w-parametric representation of the integral eq.(29). Every Bessel function $J_0(2t)$ contributes asymptotically a trigonometric function, $\cos(2t - \pi/4)$. In addition there exists another energy dependent trigonometric function, $\cos(Et)$. If we represent all these functions via complex exponentials, we obtain under the integral in eq.(29) asymptotically a product of the power function and a sum (with well determined coefficients) of exponents $\exp[i(w - w_j)t]$. These are the resonance values w_j. In the 2-D case we have a strong resonance (the function $R(w_j, E)$ diverges).

It is easy to establish that the function $R(w,E)$ satisfies eq.(A.3) for the real value of the parameter w under the condition

$$w \geq u_0 = \max\{w_j\}, \qquad \text{(A.8)}$$

where for the 2-D system $u_0 = 2 + |E|$. I.e. here exists a line of poles which starts from the point $u_0 = 2 + |E|$ (it corresponds to the value $\sigma = 0$, because here $R(u_0, E)^{-1/2} = 0$) and moves on with increasing value of the parameter σ always along the real axis. It is also easy to find that in this case no other pole lines exist. This result can be generalized for higher dimensions. So one can establish that the value of u_0 found from eq.(34) corresponds precisely to the condition eq.(A.8): we always have to deal with the same resonance.

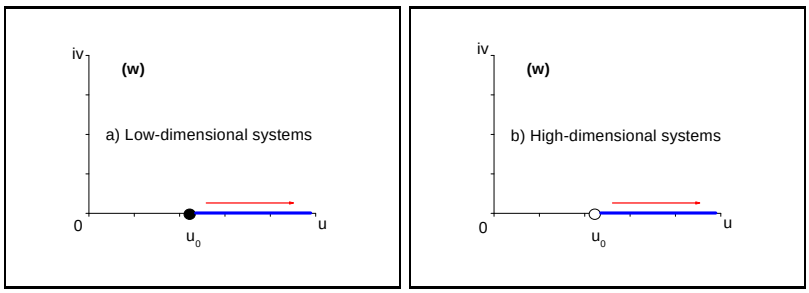

Figure A1. Parametric representation of the pole diagram for unstable filter: a) low-dimensional systems with $D = 2, 3$; b) high-dimensional systems with $D \geq 4$.

The structure of the pole diagram is nearly identical for all dimensions D, Fig.A1. There exists only one pole line, which originates at the point $w = u_0$. The direction of increasing σ is denoted by an arrow. The only difference is that for low-dimensional systems the starting point u_0 corresponds exactly to $\sigma = 0$ (this is marked in the diagram with a black circle), for high-dimensional systems the value $\sigma_0(E) = R(u_0, E)^{-1/2}$ is finite (the point $w = u_0$ is marked with a white circle). In the latter case it is impossible to find for smaller values of the disorder $\sigma < \sigma_0(E)$ a point in the diagram, where the condition eq.(A.2) is fulfilled. The filter $H_+(z)$ possesses no poles. The border thus found via eq.(37) has been defined as mobility edge.

Filter $H_-(z)$

This filter has a physical interpretation only under the condition [8] that the function $H_-(z)$ has either no poles or that they belong to the unit circle, $|z| = 1$. In the parametric w-representation the unit circle (in the sector $w = u + iv$ with $u \geq 0$, $v \geq 0$) corresponds to an interval on the real axis $u \in [0, 2]$. I.e. it is necessary first to find the single points or even lines in this interval $u \in [0, 2]$, where the condition eq.(A.3) holds. Later on one must also find out, in which way the pole line leaves this interval.

It is easy to ascertain that one of the possible pole lines always lies on the imaginary axis v, and in the interval $v \geq 0$. On this line (trivial pole line) only the point $v = 0$ could have a physical interpretation, because the point, $w = 0$, also belongs to the unit circle. The existence of other pole lines strongly depends on the space dimension D and the energy E.

As the simplest example let us first consider the filter for the 2-D system [8], where the corresponding function has an analytical form:

$$R(w, E) = \frac{1}{2\sqrt{4 - w^2}}\left[\frac{1}{\sqrt{4 - (w + E)^2}} + \frac{1}{\sqrt{4 - (w - E)^2}}\right]. \qquad \text{(A.9)}$$

Because the interval $u \in [0, 2]$ plays an important role for the physical interpretation, we consider first the resonance values $\{w_j\}$, which lie precisely in this region. It is easy to calculate that for each energy value only one resonance $w = u_0$ exists. For energies $0 \leq |E| < 2$ we have $u_0 = 2 - |E|$. In the remaining energy region, $2 < |E| < 4$, one has $u_0 = |E| - 2$. In the 2-D system the resonances are so strong (the function $R(w, E)$ diverges at $w = u_0$) that these determine in a unique manner the beginning of the new pole lines and their direction, Fig.A2.

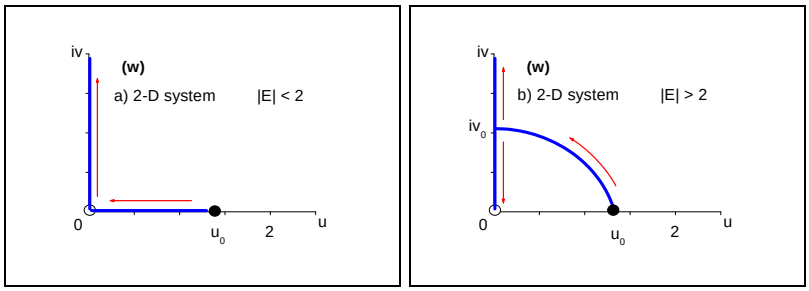

Figure A2. Parametric representation of the pole diagram for 2-D system: a) energy range $0 \leq |E| < 2$; b) energy range $2 < |E| < 4$.

In the interval $0 \leq |E| < 2$ the pole line starts exactly at the point $u_0 = 2 - |E|$. This point corresponds to the value of the disorder $\sigma = 0$ and is denoted by a black circle in the diagram. With increasing value of σ the pole line follows exactly the real axis u, until it reaches the point $u = 0$ ($w = 0$). At this point the corresponding disorder $R(0, E)^{-1/2}$ is finite (white circle in the figure), and we have denoted this in eq.(41) as $\sigma_0'(E)$. Because this line is always a unit circle (marginal stability), we have physical solutions. These exist only under the condition that $0 \leq \sigma < \sigma_0'(E)$. If the disorder crosses the border $\sigma_0'(E)$, the pole line leaves the point $w = 0$ and follows further the trivial pole line (imaginary axis). Here, however, a physical interpretation

is no longer possible and consequently there are no extended states in the range $\sigma > \sigma_0'(E)$.

In the interval $2 < |E| < 4$ the behaviour of the pole line determines the resonance at $w = u_0 = |E| - 2$. This line again starts at the value u_0, which corresponds to the disorder $\sigma = 0$ (black circle in the figure). Then this line leaves the real axis and takes its course into the complex plain until it reaches a particular point v_0 (bifurcation point) on the imaginary axis v. The later path makes use of branches of trivial pole lines. The directions belonging thereto are again marked with an arrow. Thus we have for $\sigma > 0$ no points in the phase diagram which can be interpreted physically. I.e. an infinitesimal disorder already suffices in this energy range to destroy all extended states. One may also write that in this range $\sigma_0'(E) \equiv 0$.

Formally among the 3-D case (low-dimensional system) and the cases $D \geq 4$ (high-dimensional systems) there is a quantitative difference. For high-dimensional systems one does not find a divergence of the function $R(w, E)$ at the resonance $w = u_0$, but only a jump in the argument of this complex function. For the 3-D system this divergence exists, on the other hand, but in contrast to the 2-D system the resonance is very weak (logarithmic divergence). In all these cases the resonances have no direct influence: they do not generate a pole line emerging from the point u_0. There is, however, an indirect influence of the resonances, because every resonance defines the argument of the function $R(w, E)$.

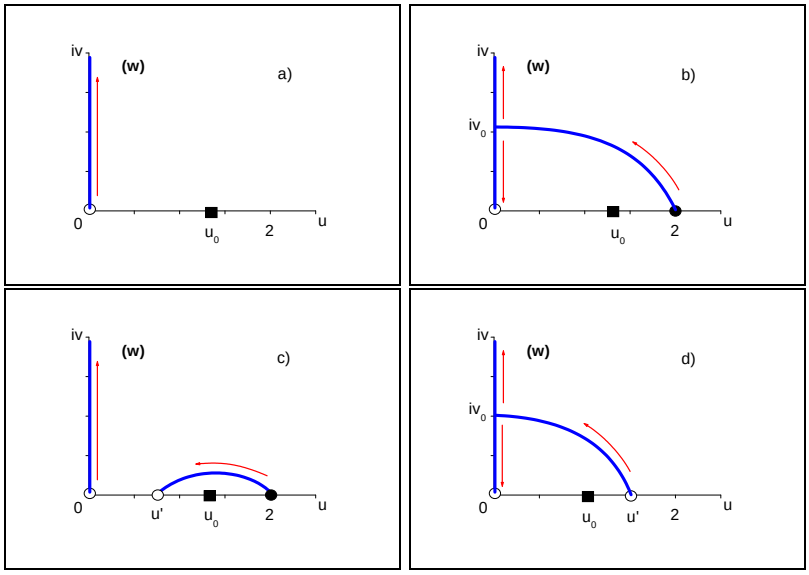

Figure A3. Different types of pole diagrams for the systems with spatial dimensions $D \geq 3$. For details see the text.

In fig.A3, different types of pole diagrams are presented. Fig.A3a is typical for small values of the energy. Although the resonance at $w = u_0$ exists (this point is marked in the figure with a black square), the weak resonance does not generate a new pole line. There remains only the trivial pole line, where only the point $w = 0$ has physical significance. The value $R(0, E)^{-1/2}$ defines again the function $\sigma_0'(E)$ already mentioned. Physical solutions exist only for $\sigma < \sigma_0'(E)$ (no poles on the unit circle).

Fig.A3b is typical for large values of the energy. In the parametric representation of the filter function, eq.(32), one can in addition see a point, where the function $R(w, E)$ diverges. This is the point $w = 2$ (a type of energy independent resonance, which corresponds to the factor $\sqrt{4 - w^2}$). One can see that this point takes over the role of the resonance at u_0. The pole line starts from the point $w = 2$ (this point corresponds to the value $\sigma = 0$ and is therefore again marked with a black circle). The continuation is similar to Fig.A2b: the line continues to the imaginary axis and reaches it at the point $v = v_0$. Although the points $w = u_0$ and $w = 2$ are not identical, Fig.A2b and Fig.A3b are qualitatively similar, and they have the same physical interpretation. In this range of the energy an infinitesimal disorder destroys all extended states.

An indirect influence by the resonances consists in the fact that one can change in their neighbourhood the argument of the complex function $R(w, E)$ in such a way that the condition eq.(A.3) becomes valid at a point $w = u'$, see Fig.A3c and Fig.A3d. The point u' is not a resonance, here the function $R(u', E)$ remains finite (white circle in the pole diagram).

Let us consider further on two examples. In the 3-D case on can divide the energy values into three ranges. In the range $0 \leq |E| < 2$ the resonance is found at $u_0 = |E|$, which, however, has no importance. We have the pole diagram of Fig.A3a. In the range $2 \leq |E| < 4$ the resonance occurs at $u_0 = 4 - |E|$. Here the above mentioned point u' arises, but under the condition $|E| > E_0 = 3.00$. I.e. in the range $2 \leq |E| < E_0$ the pole diagram of Fig.A3a remains valid. In the range $E_0 < |E| < 4$ we find a different typ of diagram, Fig.A3c. The point u' corresponds to the condition $0 \leq u' < u_0$. For $E \to 4$ the point u' moves towards the value $u' = 0$, and thus arises the bridge between the new pole line and the trivial pole line. After this the diagram is qualitatively the same as Fig.A3b: there are no extended states. This is also valid in the range $4 < |E| < 6$, where $u_0 = |E| - 4$.

Figure A3c describes a complicated case which does not have an unambiguous interpretation. Formally this is the only diagram which has two pole lines. The other ones consisted always of a single pole line, although bifurcation points were also possible. For infinitesimally small disorder the resonance at $w = 2$ is important, as in the diagram of Fig.A3b. Here with increasing disorder σ, the pole line which starts at the point $w = 2$ (black circle), leaves the interval $u \in [0, 2]$. The corresponding poles have no physical interpretation, which corresponds to the annihilation of extended states via infinitesimal disorder. If the disorder increases further this pole line approaches again the interval $u \in [0, 2]$, and reaches it at the point u' (Fig. A3c). The corresponding value of the disorder $\sigma_1 = R(u', E)^{-1/2}$ is finite (white circle). The physical interpretation now depends to which value of the disorder $\sigma_2 = R(0, E)^{-1/2}$ corresponds the point $w = 0$ (white circle) on the trivial pole line.

If $\sigma_1 < \sigma_2$, then a sort of gap arises in the disorder in such a manner that in the range $\sigma_1 < \sigma < \sigma_2$ no values of the parameter w correspond to the pole line. I.e. the filter $H(z)$ has no poles in this range. Consequently a physical interpretation of the solution is possible here and we obtain the reappearance of extended states at finite disorder values in this special interval. This condition, $\sigma_1 < \sigma_2$, is valid for 3-D systems only in a very narrow range of the energy, $E_0 < |E| < E_0'$, where $E_0' = 3.367$. The behaviour of the curve $\sigma_0'(E)$ is shown in Fig.2 (so-called reentrant behaviour). We can clearly see that to each energy value in the range $E_0 < |E| < E_0'$ are associated three values of $\sigma_0'(E)$; these are in particular $\sigma_0'(E) = 0$ - localization via infinitesimal disorder, $\sigma_0'(E) = \sigma_1$ - reappearance of extended states, and $\sigma_0'(E) = \sigma_2$ - again localization.

Outside this range, $E_0' < |E| < 4$, one has $\sigma_1 > \sigma_2$. For this condition there are no points on the pole line which permit a physical interpretation. I.e. after this annihilation of extended states via infinitesimal disorder it is not possible for extended states to reappear.

In the 4-D case we find a different sequence of resonances. In the range $0 \leq |E| < 2$ the resonance occurs at $u_0 = 2 - |E|$, and the diagram Fig.A3a is valid. In the range $2 < |E| < 4$ the resonance is found at $u_0 = |E| - 2$. Here the point u' ($u_0 > u' \geq 2$) arises. The pole diagram corresponds to A3d. The pole line starts from a white circle, which corresponds to the value $\sigma_0'(E) > 0$. Only this point in the pole diagram has a physical interpretation, and it means the existence of extended states under the condition of small values of the disorder parameter, $\sigma < \sigma_0'(E)$. For $E \to 4$ the values u' and u_0 move towards $w = 2$. Because always $u' > u_0$, this means that the point u' reaches the value $w = 2$ before the point u_0, and this occurs for $E_0 = 3.915 < 4$. For $E > E_0$ the diagram A3b is still valid, although the series of resonances is not yet exhausted: a resonance arises at $u_0 = 6 - |E|$ for $4 < |E| < 6$, and another one at $u_0 = |E| - 6$ for $6 < |E| < 8$. Here also extended states are not possible.

Appendix B. Discussion

Exact solution and scaling theory of localization

The scaling theory of Anderson localization uses the conductance g as the order parameter. It is supposed that the system size dependence of the conductance is determined only by the value of the conductance itself [2]. This relation contains no information about the microscopic structure of the model, and the tight-binding Hamiltonian is not used in the derivation of this relation. However, one should not forget that the scaling theory of localization constitutes a typical *phenomenological* theory. There is no reason to believe that a phenomenological theory correctly reproduces all critical properties of a microscopical model (in our case the tight-binding Hamiltonian). Between the scaling theory and the tight-binding model there might exist a similar relation as e.g. between the phenomenological Landau theory of phase transitions and the microscopic Ising model [18, 24]. In the best case a phenomenological theory is in the position to provide a qualitative description of the phenomena. Also qualitative differences are possible. Quantitative differences always exist and are unavoidable. In the case of the scaling theory of localization it is quite possible that the assumption that the function in the scaling relation is an analytic function of the conductance plays a similarly critical role as the corresponding assumption with respect to analytical properties of the thermodynamic potentials in the Landau theory. The role played by exactly solvable models in the theory of phase transitions [24] clearly demonstrates that exact results for microscopic models can never be replaced by phenomenological theories. The aim of the investigation of the tight-binding Hamiltonian is in no case only a confirmation of the results of scaling theory, but much more the search for possible deviations and problematic situations.

Reports of metallic behaviour in dilute two-dimensional electron-hole systems [4] render it mandatory to *reexamine* [17] the basic methods which have been used in the past years, in conjunction with the scaling theory of localization. Exact diagonalization and level statistics for finite, $L \times L$, two-dimensional systems [25] and transfer-matrix approach for strips of $\infty \times L$ [17] have proven in a numerical way the failure of single-parameter scaling in Anderson localization, at least in the

statistical distribution of electronic wave-function amplitudes.

Exact solution and numeric methods

The Anderson localization problem constitutes a multidisciplinary problem. One is not dealing with the purely quantum mechanical consequences of disorder in solids. Formally the Schrödinger equation with random on-site potentials in the tight-binding representation is a stochastic algebraic equation, where physical meaning can only be attributed to certain average values. Averaging over random potentials forces us to consider statistical ensembles of macroscopically different systems. In this sense the problem is very similar to statistical physics, especially to the statistical physics of phase transitions, because there also a metal-insulator phase transition is analyzed. In the case of phase transitions one finds that the relevant parameters (in our case e.g. Lyapunov exponents or localization length) are not analytical functions of the disorder. These non-analytical functions originate from two steps: averaging over random potentials and taking the thermodynamic limit $L \to \infty$.

The confirmation of the validity of the scaling theory of Anderson localization derives mainly from numerical studies. Whether the numerical work is in the position to simultaneously take into account all aspects of the problem can be doubted and is in itself the topic for a discussion (see below). The basic numerical methods to study the Anderson localization problem are in no case logically closed and consistent schemes, which permit a totally independent confirmation of the existence of a phase transition. One is dealing with a type of computer experiment, where one only tests particular hypotheses. If the set of such hypotheses is not complete, there is no reason to believe that the numerical results can be interpreted unequivocally. E.g. finite-size scaling theory [26] is only developed for phase transitions of second order. Finite-size scaling theory is by itself not in the position to determine independently the order of the phase transition. It only checks whether the scaling function is typical for phase transitions of second order. In the latter case it is possible to obtain via the scaling function critical values and exponents, i.e. the physical interpretation of the single-parameter scaling function in [12, 27] is correct, if the phase transition is really of second order. If the corresponding mathematics sees no phase transition of second order (and this is the case in two-dimensions), then this is no clear proof for the nonexistence of the phase transition.

For two-dimensions there exists a typical single-parameter scaling function, whose behaviour one commonly interprets as complete localization [12, 27]. We have shown in [8] that for phase transitions of first order the identical single-parameter scaling function as in [12, 27] has a totally different physical interpretation, it only describes the behaviour of the insulating phase. I.e. in this case not the numbers which the numerical works provides should be doubted but their physical interpretation. Numerical scaling is not capable of analysing a system consisting of two phases. Scaling theory of localization predicts instead that the metal-insulator transition is a continuous one in 3-D, and that all states are localized in 2-D. In this way a logical circle has been constructed. Scaling theory of localization claims that for the Anderson model only phase transitions of second order are possible or none. Finite-size scaling theory only checks this idea and interprets any deviation from a phase transition of second order as the nonexistence of the phase transition.

Statistic at the critical point

Measurements of coherent transport in mesoscopic disordered systems showed large statistical fluctuations and the non-self-averaging nature of the transport coefficients such as the conductance, g. This phenomenon is referred to as universal conductance fluctuation [28]. The physical quantities are broadly distributed. It is generally accepted that a description of this system requires distribution functions of the respective quantities [29, 30].

In the article [27] one has assumed that the full averaging ($N \to \infty$) over different realizations of disorder can be replaced by one realization of disorder ($N = 1$). Today one is not quite sure about this any more and one does an averaging for $N \sim 10$, e.g. $N = 5$ in [13]. I.e. there exists the claim that a very small number of realizations shows the full properties of the complete statistics. This claim can obviously not be proved and detailed numerical investigations [16] show clearly that the averaged values always exhibit a certain distribution.

Verification of the scaling theory of localization via the conductance distribution requires a large statistical ensemble N. In practice, however, the quality of the statistics is determined by the value of the product NL^D, which in numerical investigations is in its magnitude a constant depending on the computer. We should not forget in this context that we are dealing here with the statistics corresponding to a phase transition. It is well-known [19], that if one replaces in the statistics of a phase transition a statistical sum by the *maximal* term, then this procedure corresponds to an approximation - the mean field approximation. Because mean field theories are only correct for higher dimensions $D \geq D_0$, the results look qualitatively as if one would have artificially increased the dimension D. The replacement of the statistical sum by a small number of rather arbitrary terms is meaningless and produces no physical results. The best one can say about this numerical method in the Anderson localization problem is that it possibly considers some mean field model.

Fluctuation and first order phase transition

This discussion, however, is not fully sufficient. It does not suffice to assume that a probability distribution exists. It is more important to find out how this probability distribution arises physically. Only then can one detect, what physical importance this probability distribution possesses. Universal conductance fluctuation is really a fingerprint of the critical phenomenon, but does not fit into the generally accepted picture that the localization-delocalization transition is of second order. Phase transitions of second order do not require a description via a probability distribution, although the fluctuations at critical point are very strong and their role is well understood. If the physical origin of such fluctuations remains uncertain, then their formal description is also uncertain. There exists for the present case an experimental result which can play a decisive role. Ilani et al.[6] have studied via direct electrostatic probing the spatial structure at the metal-insulator transition in two dimensions. They found a coexistence of localized and metallic regions associated with 2-D MIT. Optical investigations suggest that the 2-D electron system becomes strongly inhomogeneous: coexistence of two lines in photoluminescence spectra, one of which is caused from metallic regions and the other proves the existence of insulating islands in the electron system [31].

From the point of view of the theory of second order phase transitions this

phenomenon is quite ununderstandable, but it has a very simple explanation in the framework of the theory of first order phase transitions. We are dealing here with a phase coexistence, where the relative proportions of the two phases varies. Then the universal conductance fluctuations are nothing else but a direct consequence of the existence of fluctuations in the heterogeneous phase. Because the transport properties are connected with percolation and their possibility depends strongly on the proportions of the two phases and their respective spatial distributions, it is clear that for different realizations of a static disorder potential the conductance of an otherwise identical mesoscopic conductor will significantly differ. The phenomenon of phase separation which originates from the existence of a first order phase transition in the 2-D electron system was discussed in [32]. It is certain that in the case of a phase transition of first order the description of heterogeneous fluctuations via probability distributions is formally possible. Whether these distributions have a clear physical interpretation is another question.

This is actually a weak point in the finite-size scaling procedure [12, 27], because we have shown in [8] that in systems with a phase transition of first order mean values are always not self-averaging quantities. True self-averaging quantities are not only those which do not fluctuate within the statistical ensemble in the thermodynamic limit, but they should also have a physical meaning, and this is not the case for first order phase transitions. In this case a formal averaging over the statistical ensemble also takes into consideration an averaging over the phases, and the resulting averages have no physical meaning.

References

[1] P.W. Anderson, Phys. Rev. **109**, 1492 (1958).
[2] E. Abrahams, P.W. Anderson, D.C. Licciardello, and T.V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).
[3] L.Molinari. J.Phys.A: Math. Gen. **25**, 513 (1992).
[4] E. Abrahams, S.V. Kravchenko, M.P. Sarachik, Rev. Mod. Phys. **73**, 251 (2001).
[5] S.V. Kravchenko, D. Simonian, M.P. Sarachik, W. Mason, and J.E. Furneaux, Phys. Rev. Lett. **77**, 4938 (1996).
[6] S. Ilani, A. Yacoby, D. Mahalu and H. Shtrikman. Science **292**, 1354 (2001)
[7] S. Ilani, A. Yacoby, D. Mahalu and H. Shtrikman, Phys. Rev. Lett. **84**, 3133 (2000)
[8] V.N. Kuzovkov, W. von Niessen, V. Kashcheyevs and O. Hein, J. Phys.: Condens. Matter, **14**, 13777 (2002).
[9] N.F. Mott, in *Electronics and Structural Properties of Amorphous Semiconductors*, edited by P.G. Le Comber and J. Mort (Academic, London, 1973), p.1.
[10] P.A. Lee and T.V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985).
[11] T.F.Weiss. Signals and systems. Lecture notes. http://umech.mit.edu/weiss/lectures.html
[12] B. Kramer and A. MacKinnon, Rep. Prog. Phys. **56**, 1469 (1993).
[13] H. Grussbach and M. Schreiber, Phys. Rev. B, **51**, 663 (1995).
[14] T.M. Rice, Nature **389**, 916 (1997).
[15] H. Kunz and B. Souillard, J.Phys. (Paris) Lett. **44**, 503 (1983).
[16] I. Travěnec and P. Markoš, Phys. Rev. B, **65**, 113109 (2002).
[17] S.L.A. de Queiroz, Phys. Rev. B, **66**, 195113 (2002).
[18] H.E. Stanley, Introduction to Phase Transition and Critical Phenomena (Oxford Univ. Press, New York, 1971).
[19] S. Ma, *Modern Theory of Critical Phenomena* (Benjamin, London, 1976).
[20] V. Kuzovkov and E. Kotomin, Rep. Prog. Phys. **51**, 1479 (1988).
[21] E. A. Kotomin and V. N. Kuzovkov, *Modern Aspects of Diffusion-Controlled Reactions: Cooperative Phenomena in Bimolecular Processes*, Vol. 34 of *Comprehensive Chemical Kinetics* (Elsevier, North Holland, Amsterdam, 1996).
[22] D. Stauffer and A. Aharony *Introduction to Percolation Theory* (Taylor and Francis, London, 1992).

[23] F.A.B.F. de Moura and M.L. Lyra, Phys.Rev. Lett., **81**, 3735 (1998).
[24] R.J.Baxter, Exactly Solved Models in Statistical Mechanics (Academic Press, London, New York, 1982).
[25] J.W. Kantelhardt and A. Bunde, Phys. Rev. B, **66**, 035118 (2002).
[26] M.N. Barber, in *Phase Transition and Critical Phenomena*, edited by C. Domb and J.L Lebowitz (Academic, New York, 1983), vol.8.
[27] A. MacKinnon and B. Kramer, Phys.Rev. Lett., **21**, 1546 (1981).
[28] M. Janssen, Phys. Rep. **295**, 2 (1998).
[29] P.W. Anderson, D.J. Thouless, E. Abrahams, and D.S Fisher, Phys. Rev. Lett. **22**, 3519 (1980).
[30] B. Shapiro, Phys. Rev. B. **34**, 4394 (1986).
[31] A.A. Shashkin, V.T. Dollgopolov, G.V. Kravchenko, M. Wendel, R. Schuster, J.P. Kotthaus, P.J. Haug, K. von Klitzing, K. Ploog, N. Nickel, and W. Schlapp, Phys. Rev. Lett. **73**, 3141 (1994).
[32] B.Spivak, Phys. Rev. B. **67**, 125205 (2003).